Evaluating Deep Learning Recommendation Model Training Scalability with the Dynamic Opera Network

Connor Imes, Andrew Rittenbach, Peng Xie, Dong In D. Kang, John Paul Walters, Stephen P. Crago
{cimes,arittenb,pengxie,dkang,jwalters,crago}@isi.edu

Abstract: Training DLRMs is increasingly dominated by all-to-all and many-to-many communication patterns. The dynamic Opera network optimizes bulk data flows using direct forwarding through time-varying circuits and has been shown to be particularly useful for all-to-all traffic patterns while remaining cost-equivalent with static network topologies. We propose co-designing DLRM models with the Opera network to improve training time while matching network infrastructure cost with a traditional fat-tree topology.

Model-A (L) and Model-I (R) DLRM traffic proportions using a fat-tree network with ring all-reduce. All-to-all/many-to-many communication from embedding tables dominates at smaller scales, but all-reduce traffic increases as MLPs are replicated at larger scales.

Model-A (L) and Model-I (R) training iteration times on fat-tree, Opera, and TopoOpt networks at 16, 128 and 1024 node scales.

Model-A (L) and Model-I (R) communication and computation task completion event breakdown for fat-tree and Opera networks at 16 (Top) and 1024 (Bottom) node scales. Note that each mark is the time at which a task completed, i.e., does not denote an entire task runtime.

Conclusion: Our initial results are promising, demonstrating up to 1.79x improvement over a fat-tree network and better performance than the TopoOpt dynamically reconfigurable network.

Varying Model-A embedding table counts at 128 nodes affects all-reduce traffic patterns and training iteration time. (Not shown here) Increasing the training batch sizes also increases the many-to-many/all-to-all data exchange proportions, resulting in a relative improvement for Opera, but is generally poor for TopoOpt.